[Document]	Patent Application
[Docket Number]	11806
[Filing Date]	March 31, 2003
[Recipient]	Patent Office Administrator
[IPC]	G05D 1/02 B25J 13/08 G06T 7/00
[Inventor]	
[Address]	c/o Kabushiki Kaisha Honda Gijutsu Kenkyusho, 4-1 Chuo 1-chome, Wako-shi, Saitama
[Name]	Yoshiaki SAKAGAMI
[Inventor]	
[Address]	c/o Kabushiki Kaisha Honda Gijutsu Kenkyusho, 4-1 Chuo 1-chome, Wako-shi, Saitama
[Name]	Koji KAWABE
[Inventor]	
[Address]	c/o Kabushiki Kaisha Honda Gijutsu Kenkyusho, 4-1 Chuo 1-chome, Wako-shi, Saitama
[Name]	Nobuo HIGAKI
[Inventor]	
[Address]	c/o Kabushiki Kaisha Honda Gijutsu Kenkyusho, 4-1 Chuo 1-chome, Wako-shi, Saitama
[Name]	Naoaki SUMIDA
[Inventor]	
[Address]	c/o Kabushiki Kaisha Honda Gijutsu Kenkyusho, 4-1 Chuo 1-chome, Wako-shi, Saitama
[Name]	Youko SAITOU
[Inventor]	
[Address]	c/o Kabushiki Kaisha Honda Gijutsu Kenkyusho, 4-1 Chuo 1-chome, Wako-shi, Saitama
[Name]	Tomonobu GOTOU
[Applicant]	
[Identification Number]	000005326
[Address]	1-1 Minami-Aoyama 2-chome, Minato-ku, Tokyo
[Name]	HONDA MOTOR CO., LTD.
[Attorney]	
[Identification Number]	100089266
[Patent Attorney]	
[Name]	Yoichi OSHIMA
[Official Fee]	
[Deposit Number]	047902
[Paid Amount]	¥21,000

[List of Attached Documents]

[Document]	Specification	1
[Document]	Drawing	1
[Document]	Abstract of Disclosure	1
[General Power of Attorney]	9715829	
[Proofing Copy]	Needed	

[Document] SPECIFICATION

[Title of the Invention]

[Scope of Claims]

[Claim 1] An image transmission system for a mobile robot equipped with sound input means and image capturing means, wherein the robot comprises:

human detecting means for detecting a human based on information from the sound input means or the image capturing means;

traveling means for traveling toward the detected human;

image cut out means for cutting out an image of the detected human according to information from the image capturing means; and

image transmitting means for transmitting the image of the human to an external device.

[Claim 2] An image transmission system for a mobile robot according to claim 1, wherein the mobile robot detects a moving object from the image information obtained from the image capturing means, and detects color information of the moving object to determine that the object is a human.

[Claim 3] An image transmission system for a mobile robot according to claim 1 or 2, wherein the mobile robot determines a direction of a sound source from the sound information obtained from the sound input means.

[Claim 4] An image transmission system for a mobile robot according to any one of claim 1-3, wherein the mobile robot comprises state monitoring means for monitoring a state of the mobile robot including at least movement information, and transmits the monitored state in addition to the image.

[Claim 5] An image transmission system for a mobile robot according to any one of claims 1-4, wherein the mobile robot changes the direction of the image

capturing means based on the information on the detected human.

[Claim 6] An image transmission system for a mobile robot according to any one of claims 1-5, wherein the mobile robot calculates a distance to the detected human according to information on the detected human, and determines a target destination of movement according to the calculated distance.

[Detailed Description of the Invention]

[0001]

[Technical Field]

The present invention relates to an image transmission system for a mobile robot.

[0002]

[Prior Art]

It is known in the past to equip a robot with a camera to monitor a prescribed location or a person and transmit the obtained image data to an observer or the like (See Patent Document 1, for example). It is also known to remote control a robot from a portable terminal (See Patent Document 2, for example).

[0003]

[Patent Document 1]

Japanese patent laid open publication No. 2002-261966 (paragraphs [0035], [0073])

[Patent Document 2]

Japanese patent laid open publication No. 2002-321180 (paragraphs [0024]-[0027])

[0004]

[Tasks to be Achieved by the Invention]

However, the aforementioned conventional robots are only capable of carrying out a predetermined task in connection with a fixed location or operating under a

command sent from a remote location, and cannot operate flexibly depending on the situations.

[0005]

[Means to Achieve the Task]

In order to solve such a problem and allow the robot to autonomously move toward an object to take a picture of and transmit the taken picture, according to the present invention, a mobile robot (1) equipped with sound input means (3a) and image capturing means (2a) comprises: human detecting means (2, 3, 4, 5) for detecting a human based on information from the sound input means (3a) or the image capturing means (2a); traveling means (12a) for traveling toward the detected human; image cut out means (4) for cutting out an image of the detected human according to information from the image capturing means (2a); and image transmitting means (11) for transmitting the image of the human to an external device.

[0006]

According to such a structure, when a human to take a picture of is detected from the captured sound or image, the robot moves toward the detected human, and cuts out the image of the human for transmission to an external device. Therefore, the robot can autonomously find a person, and transmit the image of the person.

[0007]

In particular, if the mobile robot (1) detects a moving object from the image information obtained from the image capturing means (2a), and detects color information of the moving object to determine that the object is a human, a hand waving movement or the like of a person who shows interest in the robot can be detected as a moving object. If a skin color is detected in the moving object, it can be recognized as a face or hand, and thus a human can be detected. In this way, a human can be reliably

detected.

[0008]

If the mobile robot (1) determines a direction of a sound source from the sound information obtained from the sound input means (3a), even when a person only utters a voice and does not make a significant movement, the robot can determine the direction of the voice and move toward it, whereby the robot can take a picture of the object or sound source.

[0009]

If the mobile robot (1) comprises state monitoring means (6) for monitoring a state of the mobile robot (a) including at least movement information, and transmits the monitored state in addition to the image, an observer can recognize the location of the robot and can readily go to the location when he/she wants to meet the robot.

[0010]

If the mobile robot (1) changes the direction of the image capturing means (2a) based on the information on the detected human, it is possible, for example, to determine a center line of the detected human and direct the camera toward the center line. This makes it easier to cut out the human image in such a way that the human image occupies the entire frame of the transmitted image whereby the detected human can be displayed large enough even when the recipient device has a small display screen as that of a portable terminal.

[0011]

If the mobile robot (1) calculates a distance to the detected human according to information on the detected human, and determines a target destination of movement according to the calculated distance, it is possible to move the robot to an optimum position with respect to the human to take a picture of and this ensures that the picture

of the human is always taken with a favorable resolution.

[0012]

[Preferred Embodiments of the Invention]

In the following, the present invention will be described in detail based on concrete embodiments shown in the appended drawings.

[0013]

Figure 1 is an overall block diagram of a system embodying the present invention. The illustrated embodiment uses a mobile robot 1 that is bipedal, but the robot may not be limited to the bipedal type and the robot may be of a crawler type, for example. As shown in the drawing, the mobile robot 1 comprises an image input unit 2, a speech input unit 3, an image processing unit 4 connected to the image input unit 2 and serving as an image cutting out means, a speech recognition unit 5 connected to the speech input unit 3, a robot state monitoring unit 6 serving as a state monitoring means, a human response managing unit 7 that receives signals from the image processing unit 4, speech recognition unit 5 and robot state monitoring unit 6 and serves as a human response managing means, a map database unit 8 and face database unit 9 that are connected to the human response managing unit 7, an image transmitting unit 11 serving as an image transmitting means for transmitting image data to an external device according to the image output information from the human response managing unit 7, a movement control unit 12 and a speech generating unit 13. The image input unit 2 is connected to a pair of cameras 2a that are arranged on the right and left sides and serve as an imaging means. The speech input unit 3 is connected to a pair of microphones 3a that are arranged on the right and left sides and serve as a speech input means. The cameras, microphones, image input unit 2, speech input unit 3, image processing unit 4 and speech recognition unit 5 jointly form a human detection means.

The speech generating unit 13 is connected to a loudspeaker 13a serving as a sound emitting means. Further, the movement control unit 12 is connected to a plurality of motors 12a that are provided in various articulating parts and the like of the bipedal mobile robot.

[0014]

The output signal from the image transmitting unit 11 may consist of a radio wave signal that can be used in a public telephone lines, and in such a case, the signal can be received by a general portable terminal 14. The mobile robot 1 may be equipped with or hold an external camera 15 where the camera 15 may be directed to a desired object and the obtained image data may be forwarded to the human response managing unit 7.

[0015]

The control process for the transmission of image data by the mobile robot 1 constructed as above is described in the following with reference to the flowchart of Figure 2. First of all, the state variables of the robot detected by the robot state monitoring unit 6 are forwarded to the human response managing unit 7 in step ST1. The state variables of the mobile robot 1 may include the speed and direction of movement and charged state of the battery. Appropriate sensors that can detect such state variables are provided to the robot and the outputs of the sensors are forwarded to the robot state monitoring unit 6.

[0016]

The sound captured by the microphones 3a placed on either side of the head of the robot is forwarded to the speech input unit 3 in step ST2. In step ST3, the speech recognition unit 5 performs a speech recognition process on the sound data forwarded from the speech input unit 3 by using the direction and volume of the sound or cry as

parameters. The speech recognition unit 5 can estimate the location of the source of the sound according to the difference in the sound pressure level and arrival time of the sound between the two microphones 3a. The speech recognition unit 5 can also determine if the sound is an impact sound or speech from the rising edge portion of the sound level and recognize the contents of the speech by looking up the vocabulary that is stored in advance.

[0017]

An exemplary process of speech recognition in step ST3 is described in the following with reference to the flowchart shown in Figure 3. This control flow may be executed as a subroutine of step ST3. When a robot is addressed by a human, it can be detected as a change in the sound volume. For such a purpose, the change in the sound volume is detected in step ST21 in the flowchart. The location of the source of the sound is determined in step ST22. It can be accomplished by detecting a time difference and/or a difference in sound pressure between the sounds detected by the right and left microphones 3a. A speech recognition is carried out in step ST23. This can be accomplished by detecting specific words by using such techniques as separation of sound elements and template matching. The kinds of the speech may include "hello" and "come here". If the separated sound element when a change in the sound volume has occurred does not correspond to any of those included in the vocabulary or no match with any of the words included in the template can be found, the sound is determined as not being a speech.

[0018]

Once the speech processing subroutine has been finished, the image captured by the cameras 2a placed on either side of the front part of the head is forwarded to the image input unit 2 in step ST4. Each camera 2a may consist of a CCD camera, and the

image is digitized by a frame grabber to be forwarded to the imaging processing unit 4. The image processing unit 4 extracts a moving object in step ST5.

[0019]

An example of the process of extracting a moving object in step ST5 is described in the following with reference to Figure 4. The cameras 2a are directed to the direction of the sound source recognized by the speech recognition process. If no speech is recognized, the head is turned in either direction until a moving object such as those illustrated in Figure 4 is detected, and the moving object is then extracted. Figure 4a shows a person waving his hand who is captured within a certain viewing angle 16 of the cameras 2a. Figure 4b shows a person moving his hand back and forth to beckon somebody. In such cases, the person moving his hand is recognized as a moving object.

[0020]

The flowchart of Figure 5 illustrates an example of how this process of extracting a moving object can be carried out as a subroutine process. The distance d to the captured object is measured by using stereoscopy in step ST31. The reference points for this measurement can be found in the parts containing a largest number of edge points that are in motion. In this case, the outline of the moving object is extracted by a method of dynamic outline extraction using the edge information of the captured image, and the moving object can be detected from the difference between two frames of the captured moving image that are either consecutive to each other or spaced from each other by a number of frames.

[0021]

A region for seeking a moving object is defined within a viewing angle 16 in step ST32. For example, a processed distance region ($d \pm \Delta d$) is defined with respect to the distance d, and pixels located within this distance region are extracted. The

number of pixels are counted along each of a number of vertical axial lines that are arranged laterally at a regular pixel interval in Figure 4a, and the vertical axial line containing the largest number of pixels is defined as a center line Ca of a region for seeking a moving object. A width corresponding to a typical shoulder width of a person is computed on either side of the center line Ca, and the lateral limit of the region for seeking a moving object is defined according to the computed width. A region 17 for seeking a moving object defined as described above is indicated by dotted lines in Figure 4a.

[0022]

Characteristic features are extracted in step ST33. This process may consist of seeking a specific marking or other features by pattern matching. For instance, an insignia that can be readily recognized may be attached to the person who is expected to interact with the robot in advance so that this person may be readily tracked by seeking for the insignia. A number of patterns of hand movement such as that when a person spots the robot may be stored in the system so that a person may be identified by searching for a hand movement that matches any of the stored patters.

[0023]

The outline of the moving object is extracted in step ST34. There are a number of known methods for extracting an object (such as a moving object) from given image information. The method of dividing the region based on the clustering of the characteristic quantities of pixels, outline extracting method based on the connecting of detected edges, and dynamic outline model method (Snakes) based on the deformation of a closed curve so as to minimize a pre-defined energy are among such methods. An outline is extracted from the difference in brightness between the object and background, and a center of gravity of the moving object is computed from the positions of the

points on or inside the extracted outline of the moving object. Thereby, the direction (angle) of the moving object with respect to the reference line extending straight ahead from the robot can be obtained. The distance to the moving object is then computed once again from the distance information of each pixel of the moving object whose outline has been extracted, and the position of the moving object in the actual space is determined. When there are more than one moving object within the viewing angle 16, a corresponding number of regions are defined so that the characteristic features may be extracted from each region.

[0024]

When a moving object was not detected in step ST5, the program flow returns to step ST1. Upon completion of the subroutine for extracting a moving object, a map database stored in the map database unit 8 is looked up in step ST6 so that the existence of any restricted area may be identified in addition to determining the current location and identifying a region for image processing.

[0025]

In step ST7, a small area in an upper part of the detected moving object is assumed as a face, and color information (skin color) is extracted from this area considered to be a face. If a skin color is extracted, the location of the face is determined, and the face is extracted.

[0026]

Figure 6 is a flowchart illustrating an exemplary process of extracting a face in the form of a subroutine process. Figure 7a shows an initial screen showing the image captured by the cameras 2a. The distance is detected in step ST41. This process may be similar to that of step ST31. The outline of the moving object in the image is extracted in step ST42 similarly as the process of step ST34. In these steps ST41 and 42, the data

acquired in steps ST32 and 34 may be used.

[0027]

If an outline 18 as illustrated in Figure 7b is extracted in step ST43, the positional data (top) of the uppermost part of the outline 18 in the screen is set as a head top 18a. An area of search is defined by using the head top 18a as a reference point. The area of search is defined as an area corresponding to the size of a face that depends on the distance to the object similarly as in step ST32. The depth range is also determined by considering the size of the face.

[0028]

The skin color is then extracted in step ST44. The skin color region can be extracted by performing a thresholding process in the HLS (hue, lightness and saturation) space. The position of the face can be determined as a center of gravity of the skin color region within the search area. The processing area for a face which can be assumed to have a face size depending on the distance to the object is defined as an elliptic model 19 as shown in Figure 8.

[0029]

Eyes are extracted in step ST45 by detecting black circles (eyes) within the elliptic model 19 defined as described earlier by using a circular edge extracting filter. A black circle search area 19a having a certain width (depending on the distance to the person) is defined according to a standard position of eyes as measured from the head top 18a, and the eyes are detected easily by performing the eye searching within the area 19a.

[0030]

The face image is then cut out for transmission in step ST46. The size of the face image is preferably selected in such a manner that the face image substantially

entirely fills up the frame of the cut out image 20 as illustrated in Figure 9 when the display screen of the recipient terminal is small such as when the recipient terminal consists of a portable terminal 14, for example. Conversely, when the display of the recipient consists of a large screen, the background may also be included in the cut out image. The zooming in and out of the face image may be carried out according to the space between the two eyes that is computed from the positions of the eyes detected in step ST45. When the face image occupies the substantially entire area of the cut out image 20, the image may preferably be cut out in such a manner that the mid point between the two eyes is located at a prescribed location (for instance, slightly above the central point of the cut out image 20). The subroutine for the face extracting process is then concluded.

[0031]

The face database stored in the face database unit 9 is looked up in step ST8. When a matching face data is detected, for instance, the name included in the associated personal information is forwarded to the human response management unit 7 along with the face image itself.

[0032]

The person whose face was extracted in step ST7 is identified in step ST9. This identification process can be conducted based on pattern recognition, correspondence estimation according to principal component analysis, and facial expression recognition.

[0033]

The position of the hands of the recognized person is determined in step ST10. The position of the hand can be determined in relation with the position of the face or by searching the skin color area defined inside the outline extracted in step ST5. In other words, the outline covers the head and body of the person, and skin color areas

other than the face can be considered as hands because only the face and hands are normally exposed.

[0034]

The gesture and posture of the person are recognized in step ST11. The gesture as used herein may include particular body movements such as waving a hand and beckoning some one by moving a hand that can be detected by considering the positional relationship between the face and hand. The posture may consist of any bodily posture that indicates that the person is looking at the robot. Even when a face was not detected in step ST7, the program flow advances to this step ST11.

[0035]

A response to the detected person is made in step ST12. The response may include speaking to or moving toward the detected person and directing the camera and/or microphone toward the detected person by turning the head of the robot toward the detected person. In step ST13, the image of the detected person that has been extracted in the steps up to step ST12 is compressed for the convenience of handling, and an image converted into a format that suits the recipient of the transmission is transmitted. Preferably, the states of the mobile robot 1 detected by the robot state monitoring unit 6 may be superimposed on the image. Thereby, the position and traveling speed of the mobile robot 1 can be readily determined by simply looking at the display, and the operator of the robot can easily know the state of the robot by means of a portable remote terminal.

[0036]

By thus allowing a person to be extracted by the mobile robot 1 and the image of the person to be received by a portable remote terminal 14 via public lines, one can view the scene and person captured by the mobile robot 1 at will by using the portable

terminal 14. For instance, when a long line of people has been formed in an event hall, the robot may speak to people who are bored from waiting. The robot may move toward one of those people who showed interest in the robot, and capture the scene while chatting with the person so that this scene may be shown on a large display on the wall or the like . If the robot 1 carries a camera 15, the image acquired by the camera may be transmitted similarly as above so that the acquired image can be displayed on the monitor of a portable remote terminal 14 or a large screen on the wall.

[0037]

When a face was not detected in step ST7, the robot approaches what appears to be a human according to the gesture or posture analyzed in step ST11, and determines an object closest to the robot from those that appear to have waved a hand, for example. The image of the object is then cut out so as to fill the designated display area 20 as shown in Figure 10, and this cut out image is transmitted. In this case, the size is adjusted in such a manner that the vertical length or lateral width, whichever is greater, of the outline of the object fits into the designated area 20 for the cut out image.

[0038]

The mobile robot 1 may be used for looking after children who are separated from their parents in places such as event halls where a large number of people congregate. The control flow of an exemplary task of looking after such a separated child is shown in the flowchart of Figure 11. The overall flow may be generally based on the control flow illustrated in Figure 2, and only a part of the control flow that is specific to looking after a separated child is described in the following along the flowchart of Figure 11.

[0039]

In this process for looking after a separated child, a fixed camera takes a

picture of the face of each child at the entrance to the event hall, for example, and this image is transmitted to the mobile robot 1. The mobile robot 1 receives this image by using a wireless receiver not shown in the drawing, and the human response management unit 7 registers the face image data in the face database unit 9. If the parent of the child has a portable terminal equipped with a camera, the telephone number of this portable terminal is also registered.

[0040]

Similarly as in steps ST21 to ST23, the detection of change in the sound volume, determination of the direction of sound source, and speech recognition are performed in steps ST51 to 53. In the step ST5 [*sic*], it is preferred if the crying of a child has been registered as a special item of the vocabulary. A moving object is detected in step ST54 similarly as in step ST5. Even when a crying of a child is not detected in step ST53, the program flow advances to step ST54. Even when a moving object is not extracted in step ST54, the program flow advances to step ST55.

[0041]

Various features are extracted in step ST55 similarly as in step ST33, and an outline is extracted in step ST56 similarly as in step ST34. A face is extracted in step ST57 similarly as in step ST 7. In this manner, a series of steps from the detection of a skin color to the cutting out of a face image are executed similarly as in steps ST43 to 46. During the process of extracting an outline and a face, the height of the detected person (H in Figure 12a) is computed from the distance to the object, position of the head and direction of the camera 2a, and the person is determined to be a child if the height is considered to be that of a child (for instance when the height is less than 120 cm).

[0042]

The face database is looked up in step ST58 similarly as in step ST8, and a person is identified corresponding to any one of the faces registered in the face database in step ST59 before the control flow advances to step ST60. Even when the person cannot be identified as a registered person, the program flow advances to step ST60.

[0043]

The gesture/posture of the detected person is recognized in step ST60 similarly as in step ST11. As illustrated in Figure 12a, when the face and the palm of the hand are considered to be close to each other based on the outline and skin color information, small movements of the face and/or hand may be recognized as a gesture. A state where a part considered to be an arm of a person according to the outline information is positioned near the head but the palm of the hand cannot be detected may be recognized as a posture.

[0044]

A human response process is conducted in step ST61 similarly as in step ST12. In this case, the mobile robot 1 moves toward the person who appears to be a child separated from its parent and directs the camera toward it by turning the face of the robot toward it. The robot then speaks to the child in an appropriate fashion by using the speaker 13a. For instance, the robot may say to the child, "Are you all right ?" Particularly when the individual person was identified in step ST59, the robot may say the name of the person. The current position is then identified by looking up the map database in step ST62 similarly as in step ST6.

[0045]

The image of the separated child is cut out in step ST63 as illustrated in Figure 12b. This process can be carried out as in steps ST41 to 46. Because the clothes of the separated child may help identify it, the size of the cut out image may be selected such

that the entire torso of the child from the waist up is included in the image.

[0046]

The cut out image is then transmitted in step ST64 similarly as in step ST13. The current position information and individual identification information (name) may also be attached to the transmitted image of the separated child, as shown in Figure 12b. If the face cannot be found in the face database and the name of the separated child cannot be identified, only the current position is attached to the transmitted image. If the identity of the child can be determined and the telephone number of the remote terminal of the parent is registered, the face image may be transmitted to this remote terminal directly. Thereby, the parent can visually identify his or her child, and can meet it according to the current position information. If the identity of the child cannot be determined, the image may be shown on a large screen for the parent to see.

[0047]

[Effects of the Invention]

Thus, according to the present invention, when a human to take a picture of is detected from the captured sound or image, the robot moves toward the detected human, and cuts out the image of the human for transmission to an external device. Therefore, the robot can autonomously find a person, and transmit the image of the person. In this way, the detection of a person and transmission of image of the person can be carried out by the robot without need for commands from the operator, and this can impose less limitation to the condition where the robot can conduct image capturing of a person and thus create a wider area of usage. The image transmitted from the robot can be visually recognized by using a portable terminal or the like, and therefore, anyone having a portable terminal and allowed to access the image can see the image at will. Thus, for example, even when one cannot see a desired person or attraction from a close range in

an event hall or the like, the one can easily see the person or attraction by using the display on his/her portable terminal.

[0048]

In particular, by recognizing a human by detecting a moving object and color information (for instance, skin color), it is possible to recognize a human easily and thus achieve a less complicated program and reduced cost. Also, it is possible to easily determine the direction of the sound source by using a stereo microphones, and this allows the robot to move toward a person who only utters a voice to take a picture of the person, which can be useful in emergency rescue. Further, if the location of the robot is transmitted as movement information, a person who is interested in the image can go to the location, and this helps monitoring various places in an event hall or the like efficiently.

[0049]

If the robot can change the direction of the cameras so as to be directed to the person, it becomes easier to extract the person and cut out an image of the same so that the person can be displayed large enough even when the recipient device has a small display screen as that of a portable terminal. For example, when a child separated from its parent is found, the image of the child is transmitted such that the child image occupies an entire area of the image, and this allows the parent to easily identify their child on his/her portable terminal. By calculating the distance to the person to determine the target destination of movement, it is possible to move the robot to an optimum position with respect to the person to take a picture of and this ensures that the picture of the person is always taken with a favorable resolution.

[Brief Description of the Drawings]

[Figure 1]

An overall block diagram of the system embodying the present invention.

[Figure 2]

A flowchart showing an example of a control mode according to the present invention.

[Figure 3]

A flowchart showing an exemplary process for speech recognition.

[Figure 4]

Figure 4 (a) is a view showing an exemplary moving object, while Figure 4(b) is a view similar to Figure 4(a) showing another example a moving object.

[Figure 5]

A flowchart showing an exemplary process for outline extraction.

[Figure 6]

A flowchart showing an exemplary process for cutting out a face image.

[Figure 7]

Figure 7(a) is a view of a captured image when a human is detected, while Figure 7(b) is a view showing a human outline extracted from the captured image.

[Figure 8]

A view showing a mode of extracting the eyes from the face.

[Figure 9]

A view showing an exemplary image for transmission.

[Figure 10]

A view showing an exemplary process of recognizing a human from his or her gesture or posture.

[Figure 11]

A flowchart showing the process of detecting a child who has been separated

from its parent.

[Figure 12]

Figure 12(a) is a view showing how various characteristics are extracted from the separated child, and Figure 12(b) is a view showing an exemplary transmission image of a child separated from its parent.

[List of the Numerals]

1	mobile robot
2	image input unit (human detecting means)
2a	camera (imaging means)
3	speech input unit (human detecting means)
3a	microphone (speech input means)
4	image processing unit (human detecting means, image cut out means)
5	speech recognition unit (human detecting means)
6	robot state monitoring unit (state monitoring means)
11	image transmitting unit (image transmitting means)
12a	motor (traveling means)

[Document] ABSTRACT OF THE DISCLOSURE

[Abstract of the Disclosure]

[Object] To allow the robot to autonomously move toward an object to take a picture of and transmit the taken picture.

[Means to Achieve the Object] The robot recognizes a moving object within an image taken by a camera 2a or when it recognizes a sound with a microphone 3a, the robot directs the camera to that direction, and detect a human by detecting color information. An image of the detected human is cut out and transmitted to an external device. The robot can autonomously find a human and transmit the image thereof. In this way, the detection of a human and transmission of image thereof can be carried out by the robot without need for commands from the operator, and this can impose less limitation to the condition where the robot can conduct image taking of a person and thus create a wider area of usage. Because the transmitted image can be visually recognized by means of a portable terminal or the like, anyone having a portable terminal and allowed to access the image can see the image at will. Even when one cannot see a desired person or attraction from a close range in an event hall or the like, the one can easily see the person or attraction by using the display on his/her portable terminal.

[Designated Drawing] Figure 1

OIPE IAP42
JAN 0 7 2008
PATENT & TRADEMARK OFFICE

Fig. 1

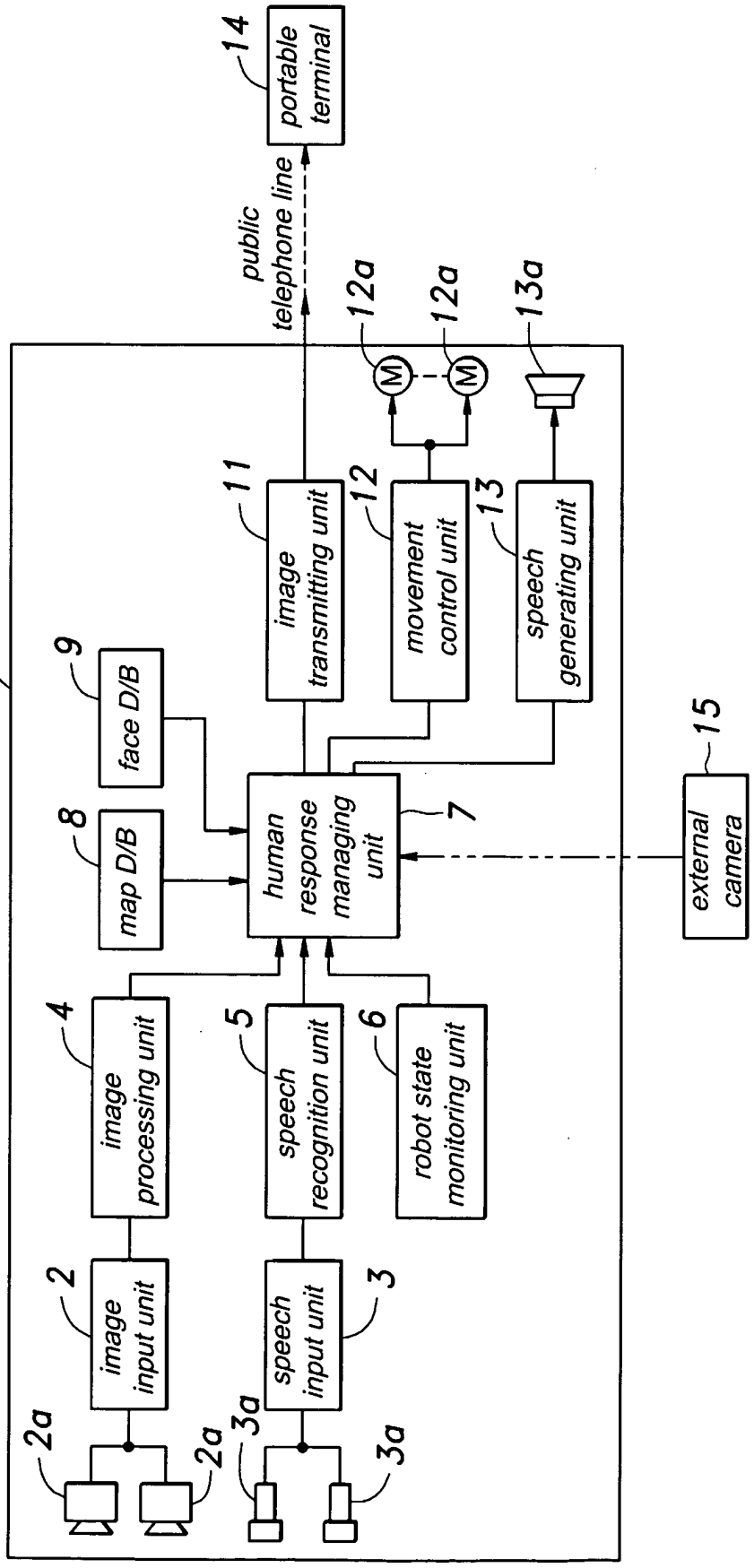

Fig.2

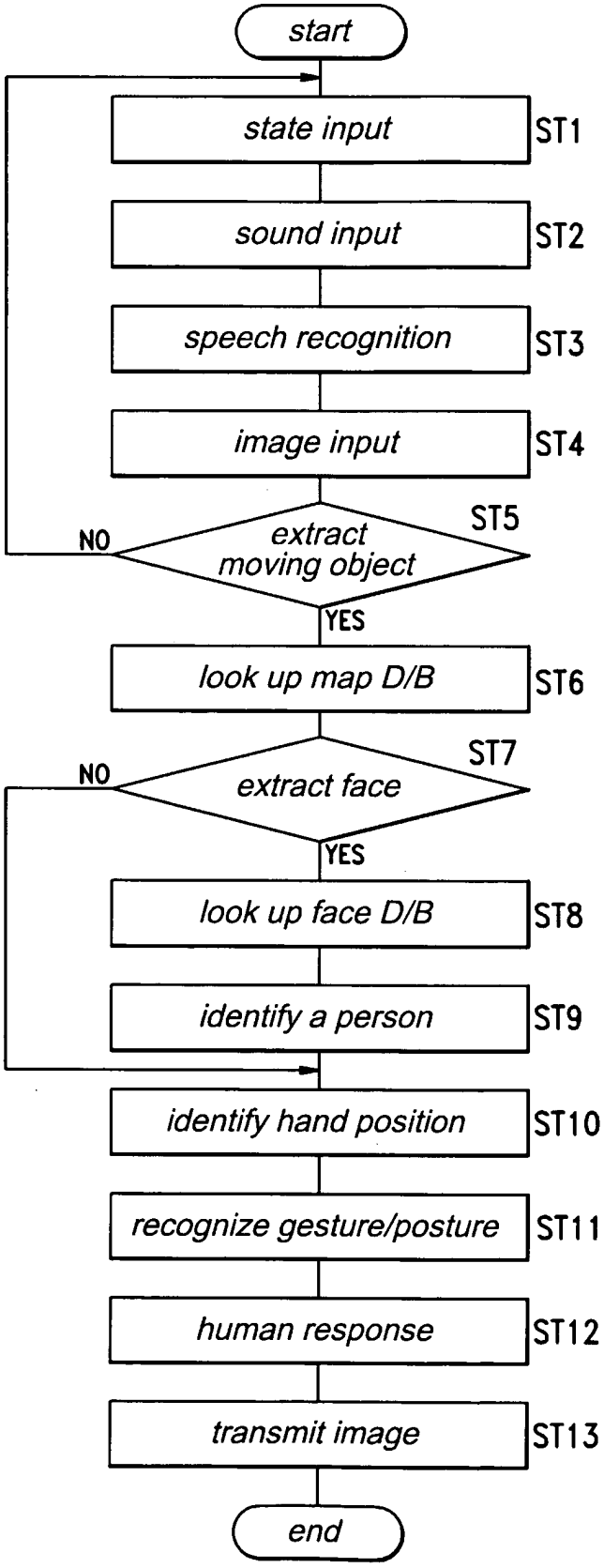
start
state input
ST1
sound input
ST2
speech recognition
ST3
image input
ST4
extract moving object
ST5
NO
YES
look up map D/B
ST6
extract face
ST7
NO
YES
look up face D/B
ST8
identify a person
ST9
identify hand position
ST10
recognize gesture/posture
ST11
human response
ST12
transmit image
ST13
end

Fig.3

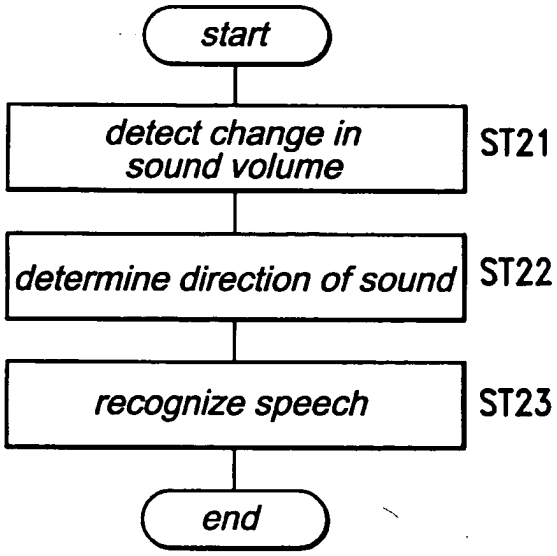

Fig.4a

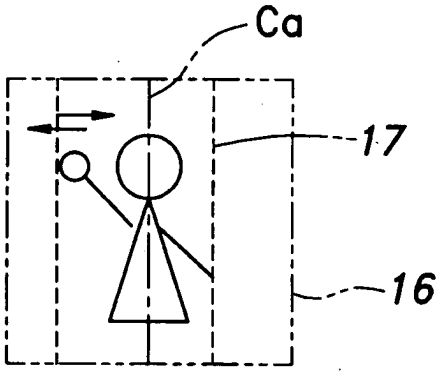

Fig.4b

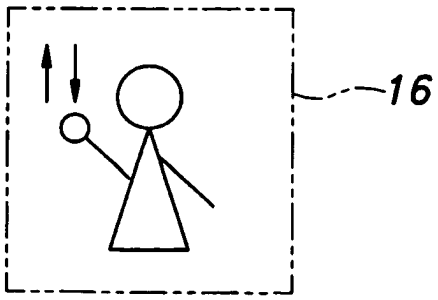

Fig.5

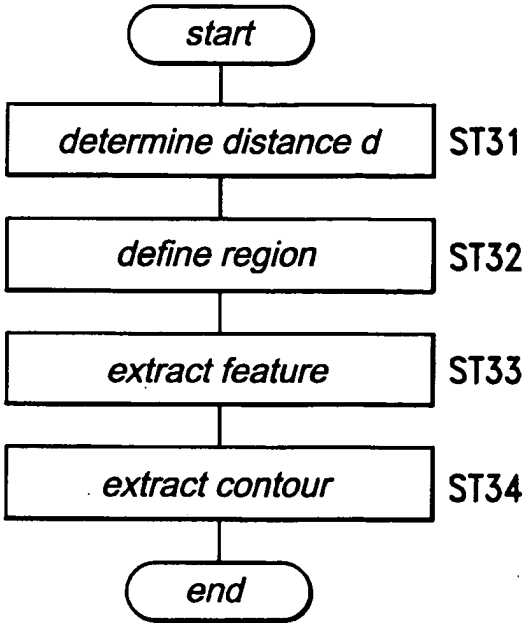
start
determine distance d
ST31
define region
ST32
extract feature
ST33
extract contour
ST34
end

Fig.6

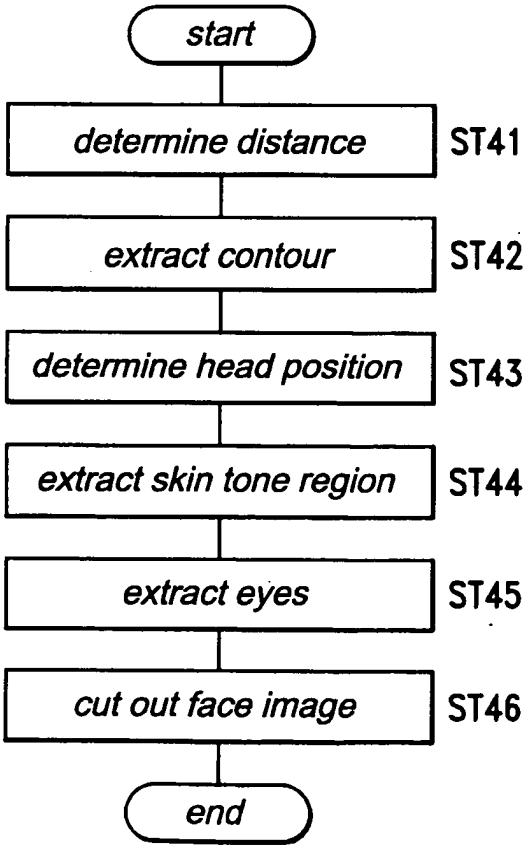
start
determine distance
ST41
extract contour
ST42
determine head position
ST43
extract skin tone region
ST44
extract eyes
ST45
cut out face image
ST46
end

Fig.7a

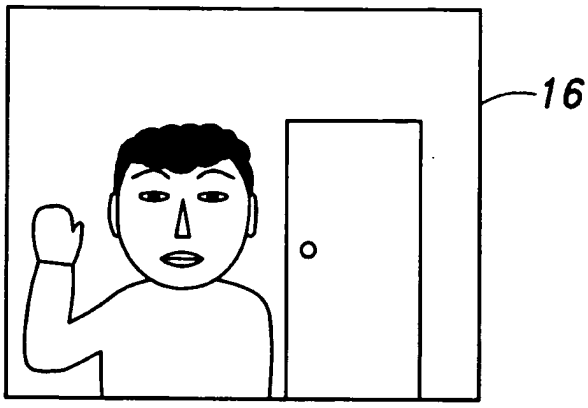

Fig.7b

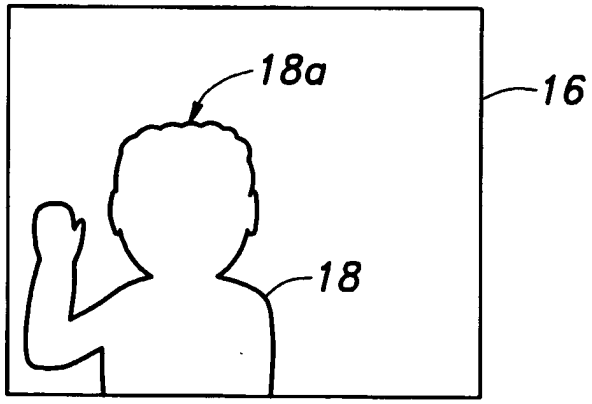

Fig.8

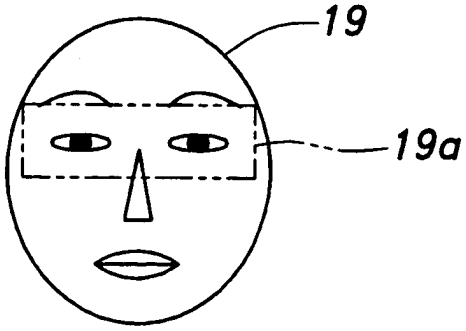

Fig.9

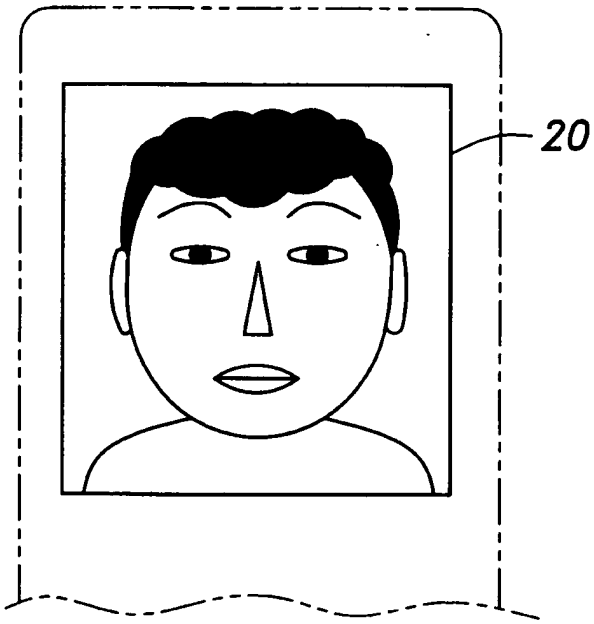

Fig.10

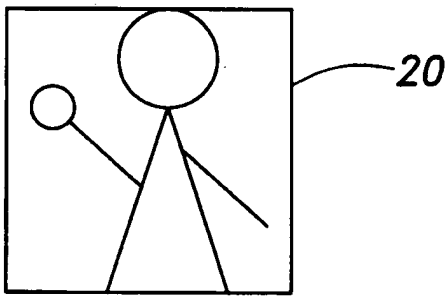

Fig.11

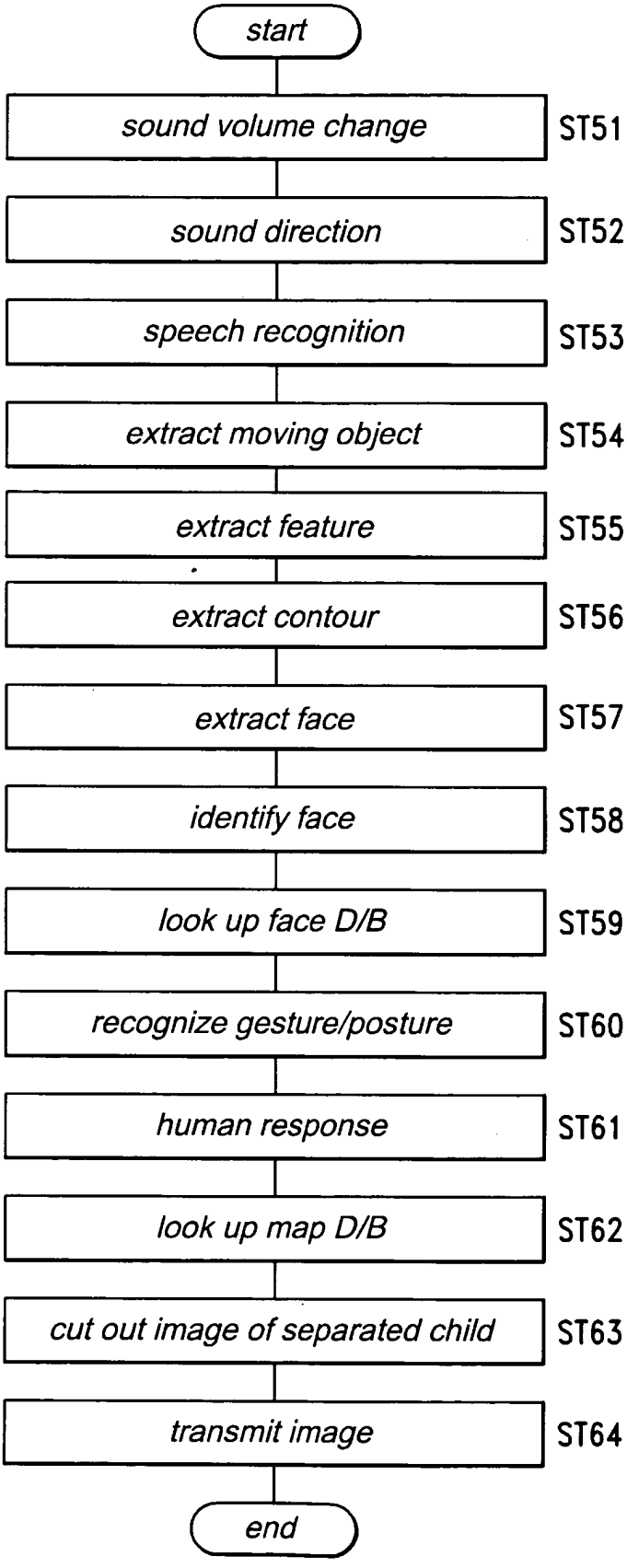
start
sound volume change
ST51
sound direction
ST52
speech recognition
ST53
extract moving object
ST54
extract feature
ST55
extract contour
ST56
extract face
ST57
identify face
ST58
look up face D/B
ST59
recognize gesture/posture
ST60
human response
ST61
look up map D/B
ST62
cut out image of separated child
ST63
transmit image
ST64
end

Fig.12a

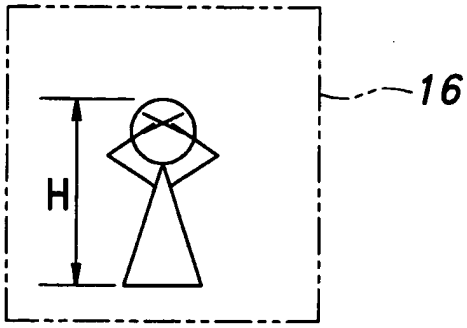
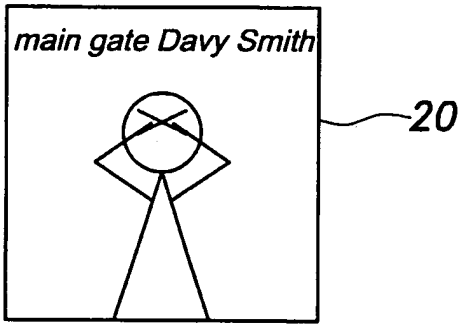
Fig.12b